

Сколько стоит разработать собственную сборку Hadoop
История и техника как получилось в Сбере
Семён Орлов
HL
HighLoad++
2022

Привет, я Семён Орлов



СБЕР

- Product Owner SDP Hadoop
- 2 ВО: физика полупроводников
  и менеджмент
- 1/3 жизни в ИТ.
  До прихода в Сбер работал в





- 1/6 жизни посвятил работе с



hdfs dfs –setrep –w 3

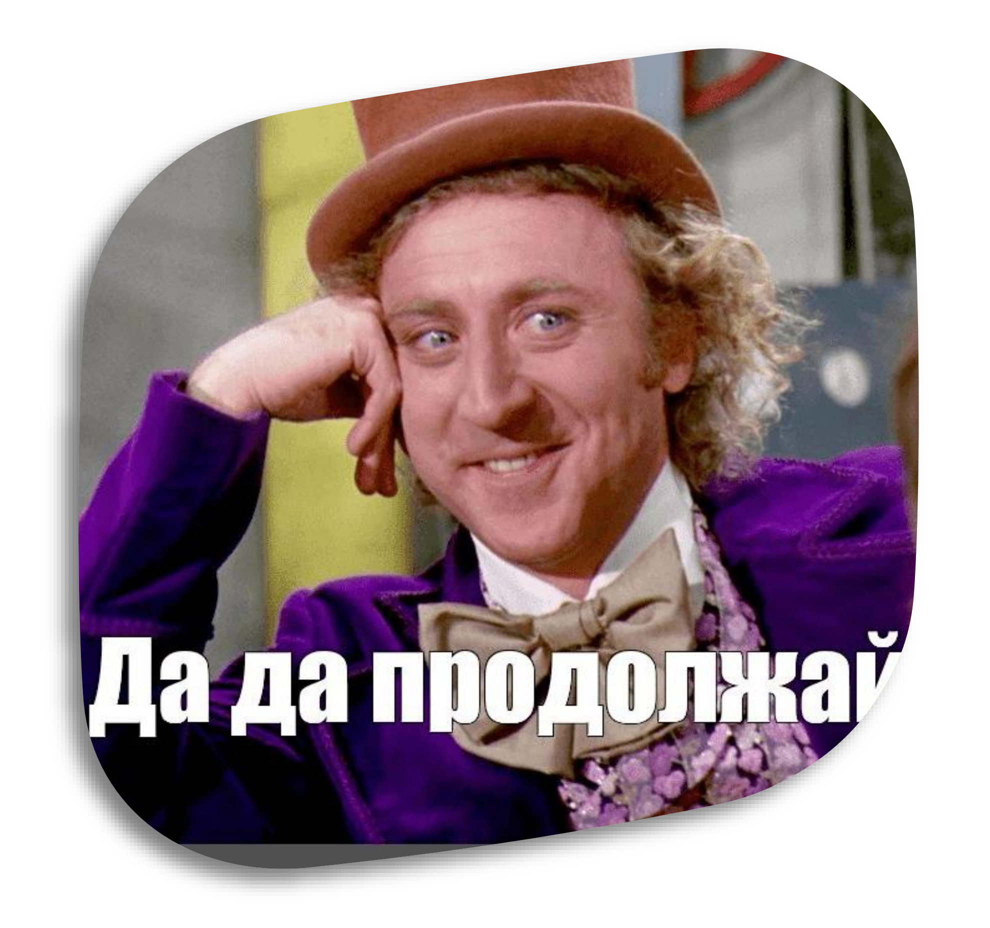


О чём расскажу?



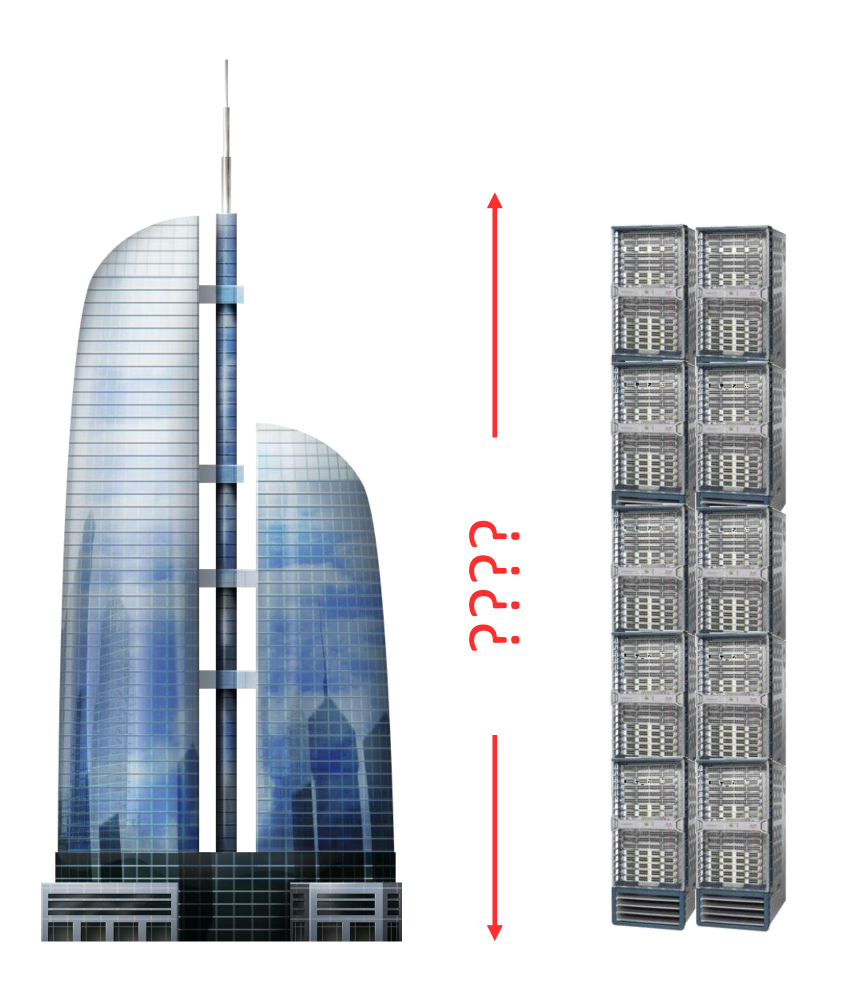
Историю, как делали

Если нужна выжимка, мотай +25 минут





Масштабы сегодня















Окружение 2018 года





1 trillion $ Company



Окружение

INHOUSE

VENDOR



Окружение лифты

СБЕР + HADOOP = SDP HADOOP



Зачем начали делать? Мотивы

Поддержка

- Сложная схема поддержки
- Невозможность работы on site
- Затраты на архитектурную поддержку

Кастом

- Безопасность
- Интеграции в КАП
- Автоматизация

TCO

- Существенная доля - лицензии
- Снижение стоимости 1ТБ

Экосистема

- Sber#
- Cloud
- Гостех



Продать идею бизнесу

1

TCO

2

Поддержка



Плюшки

- Скорость обновлений
- Компоненты
- Вендор-риск

Получить ресурсы

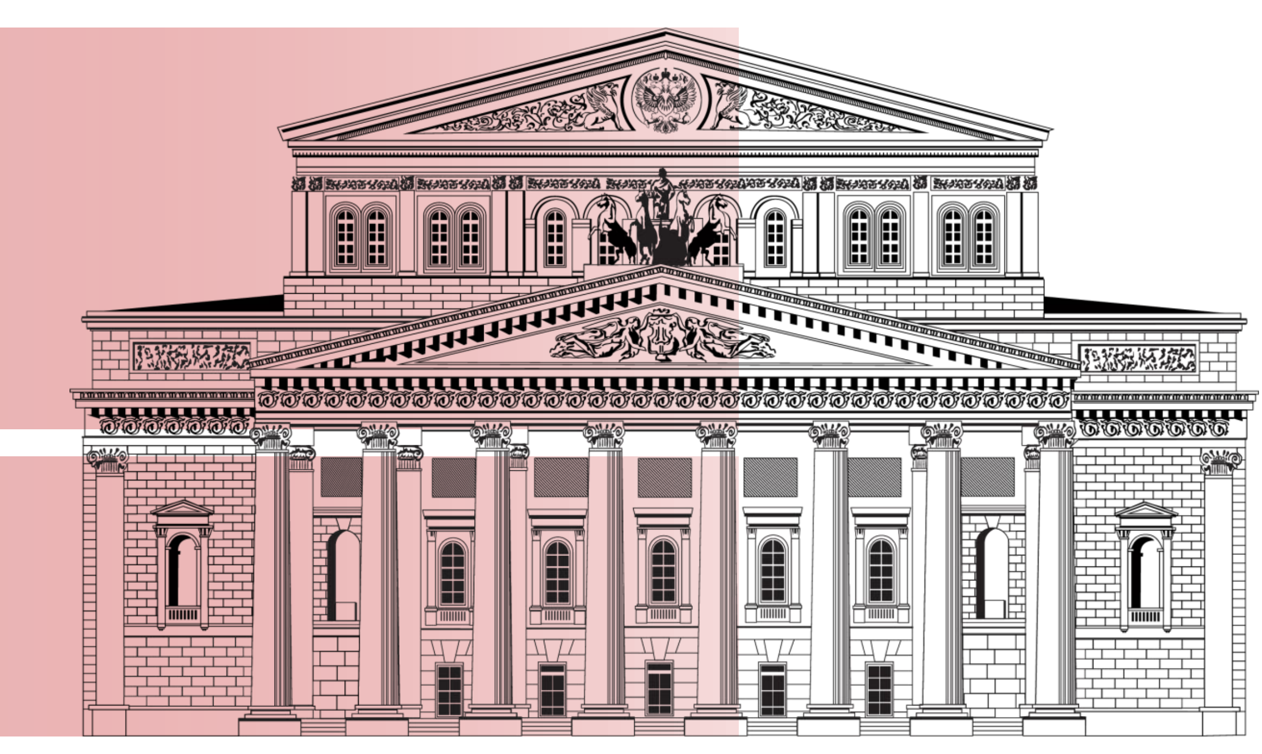


Ключ: команда для «узких» кейсов

2000 x2 **879** зрительских мест

±1 200+500+
+300 vs 10

500/36 vs 10/15

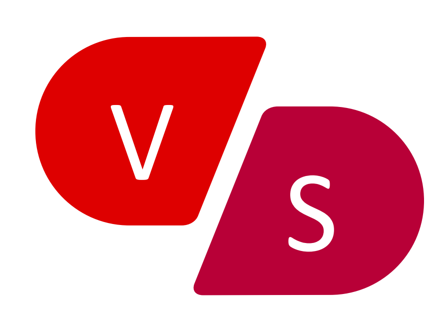




Метрики или результаты:
что важнее для выживания

Базовые метрики

- MAU/DAU
- CPI
- ARPU
- LTV
- RR



Результаты

- Кластеры в PROD
- Бесшовная миграция
- Администраторы могут сопровождать
- Безопасность не дырявая



EOL: product vision не поможет



Идея



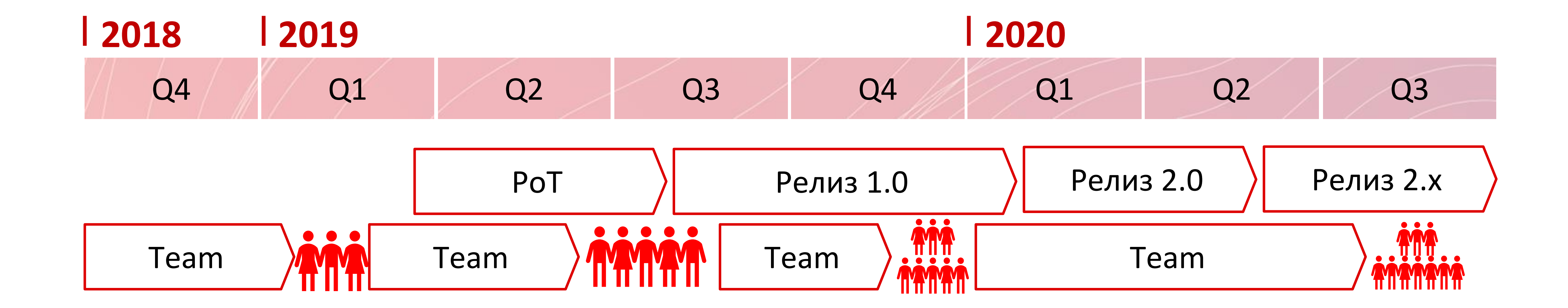
Оно работает



EOL Hadoop 2



PROD кластер



Вывод: опыт не пропьёшь, всех знаний не узнать



Собрать мало, мигрируй

2019		2020		2021		2022	
H1	H2	H1	H2	H1	H2	H1	H2

PoT → Релиз 1.0 → Релиз 2.0 → Релиз 2.x → Релиз 2.x → Релиз 2.x → Релиз 2.x → Релиз 2.x → Релиз 2.x → Релиз 2.x → Релиз 2.x → Релиз 2.x

Prepare → Миграция vendor-> SDP



Проект миграции

100+ кластеров | 1 000+ экземпляров ППО | 50 тех сервисов

Комиссия по инфраструктуре

еженедельно

- статус
- изменение планов

Экспертная площадка по разработке ППО для Hadoop3

ежемесячно

адаптациия ППО

Рабочая группа миграции

еженедельно

- статус

Статус миграции на Hadoop3

ежемесячно

- изменение планов
- эскалируем

РГ Фабрики Данных

по необходимости

- изменение планов
- эскалируем

Рабочие встречи-эскалации

по необходимости

Схема миграции

ПЯТНАШКИ
С СЕРВЕРАМИ

+20%

CDH Cluster: APP1, APP2, APP3 | SDP Cluster: APP3

CDH Cluster: APP1, APP2 | SDP Cluster: APP2, APP3

CDH Cluster: APP1 | SDP Cluster: APP1, APP2, APP3



Какие задачи на Hadoop

А что дорабатывали?

- DWH
- Реплики
  batch и NRT
  9 фреймворков
- DataMarts
- ML
- Integrations:
  a lot

Корпоративная аналитическая платформа (КАП)

Сервисы загрузки данных
- SYS603.10.01 Потоковая загрузка данных в КАП
- SYS603.20 Пакетная загрузка данных в КАП
- SYS603.20.01 Пакетная загрузка данных, Stork
- SYS603.20.02 Пакетная загрузка данных, ETL4Hadoop
- SYS603.20.03 Пакетная загрузка данных, AIR
- SYS603.10.03 Сервис загрузки данных xStream
- SYS603.10.04 Ввод пользовательских данных
- SYS603.10.05 Хеширование и токенизация при загрузке данных из контура PCI DSS
- SYS603.10.06 Загрузка данных из архивных АС
- SYS603.10.07 Хеширование при загрузке данных из Альфа
- SYS603.10.08 Шлюз ПКД в контуре PCI DSS КАП
- SYS603.10.09 Пакетная загрузка данных, ПКД
- SYS603.10.10 Оптимизация данных реплик (Data Shaper)

Приложения Ядра КАП
- КАП801 Портал Супермаркета данных
- КАП802 Первичные данные
- КАП803 Интегрированные данные (ЕСС)
- КАП804 Опубликованные данные АХД
- КАП806 Графовые данные
- КАП805 Выполнение подписок
- КАП807 Управление НСИ
- КАП808 Разделяемое хранилище
- КАП809 Геоданные
- КАП811 Архивы детальных данных АХД
- КАП812 Панель кибербезопасности
- КАП813 Карта данных
- КАП814 Контроль качества данных
- КАП815 Унифицированные клиентские данные
- КАП816 Индексированные данные
- КАП821 Портал инфоактивов
- КАП822 Описание и разметка инфоактивов
- КАП823 Управление ассортиментом Супермаркета данных
- КАП824 Аналитика и отчётность информационных активов

Сервисы обработки данных
- SYS602.26 Пакетная обработка данных в КАП
- SYS602.26.01 SberETL
- SYS602.29 Потоковая обработка данных КАП
- SYS602.29.01 Потоковая обработка на базе Apache NiFi
- SYS602.50 Инструменты BI
- SYS602.50.01 Инструменты BI - Qlik Sense
- SYS610.06 Пакетное исполнение моделей (ПИМ)
- SYS602.23 Сертификация данных КАП
- SYS602.24 Корректировки данных в КАП
- SYS602.25 Унификация клиентских данных КАП
- SYS602.48 Геообработка
- SYS602.27 Графовая платформа
- SYS602.27.01 Визуализация графовых данных
- SYS602.27.02 Графовый процессор
- SYS602.27.03 Графовые расчёты
- SYS602.27.04 Встраиваемая графовая база данных

Сервисы хранения данных
- SYS605.05.01 Хранение данных КАП в SDP
- SYS605.05.02 Хранение данных КАП в Greenplum
- SYS605.05.03 Хранение данных КАП в Teradata
- SYS605.05.04 Хранение данных КАП в Cloudera Hadoop
- SYS605.05.05 Хранение данных КАП в PostgreSQL
- SYS605.05.06 Хранение данных КАП в Oracle Exadata
- SYS605.05.07 Хранение данных КАП в Oracle DB

Сервисы управления и ИБ
- SYS604.11 Защита данных КАП
- SYS604.11.01 Сервис защиты данных КАП (Protegrity)
- SYS604.11.02 Сервис защиты данных КАП (SDG)
- SYS604.11.03 Алгоритмы обезличивания
- SYS604.09.02 Управление доступом и сертификатами КАП
- SYS609.06 Сервис управления потоками данных
- SYS604.14 Сервис безопасной выгрузки КАП
- SYS605.13 Управление метаданными КАП
- SYS601.01.03 Журналирование КАП
- SYS601.08 Мониторинг КАП
- SYS605.19 Сервисные сообщения КАП
- SYS608.18 Инструменты администрирования КАП
- SYS608.21 Сервис управления сетевыми взаимодействиями
- SYS608.20 Инструменты DataOps КАП
- SYS605.26 Анализ использования данных
- SYS608.22 Сервис управления правилами и моделями доступа

Сервис доступа к данным КАП
- SYS602.30 SQL-доступ к данным КАП
- SYS602.31 Доступ к данным КАП в реальном времени
- SYS602.32 Потоковая выгрузка данных КАП
- SYS602.33 Единый сервис поставки данных
- SYS602.33.02 Пакетная выгрузка из КАП в ТФС
- SYS602.44 Гетерогенный доступ к данным КАП
- SYS602.44.01 Гетерогенный доступ к данным КАП (QG)
- SYS602.44.02 Гетерогенный доступ к данным КАП (PXF)

Сервисы федеративного доступа
- SYS605.22.01 Сервис федеративного доступа HDFS (HDFS Gateway)
- SYS605.22.02 Сервис федеративного доступа HMS (WaggleDance)
- SYS605.22.03 Сервис федеративного доступа для авторизации Hadoop (Sentry-proxy)
- SYS605.22.04 Сервис федеративного доступа HBASE (HBASE gateway)

Аналитические приложения
- Приложения КАП
- Витрины данных
- Аналитика
- Отчётность
- Лаборатория данных
- ПКАП1001 ЛД Машина DS
- ПКАП1002 ЛД Кластер CDH
- ПКАП1003 ЛД Кластер SDP
- ПКАП1004 ЛД Лаборатория на Teradata
- ПКАП1005 ЛД Лаборатория на Greenplum
- ПКАП1006 ЛД Лаборатория на Oracle
- ПКАП1007 ЛД Лаборатория на PostgreSQL
- ПКАП1008 ЛД Лаборатория геоанализа
- ПКАП1009 ЛД Лаборатория анализа данных



Сложности  2 → 3

Cluster1
Hadoop 2

N-2
Hadoop 2

ClusterN
Hadoop 2

Hadoop
BlockIdentifier

StorageType: SSD, DISK, ARCHIVE

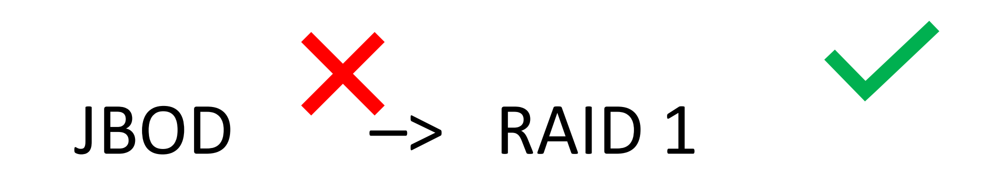
Cluster1
Hadoop 3

N-2
Hadoop 3

ClusterN
Hadoop 3



Федерации
100х кластеров



Information Lifecycle Management

Proxy, Federation

Доработки кибер-безы

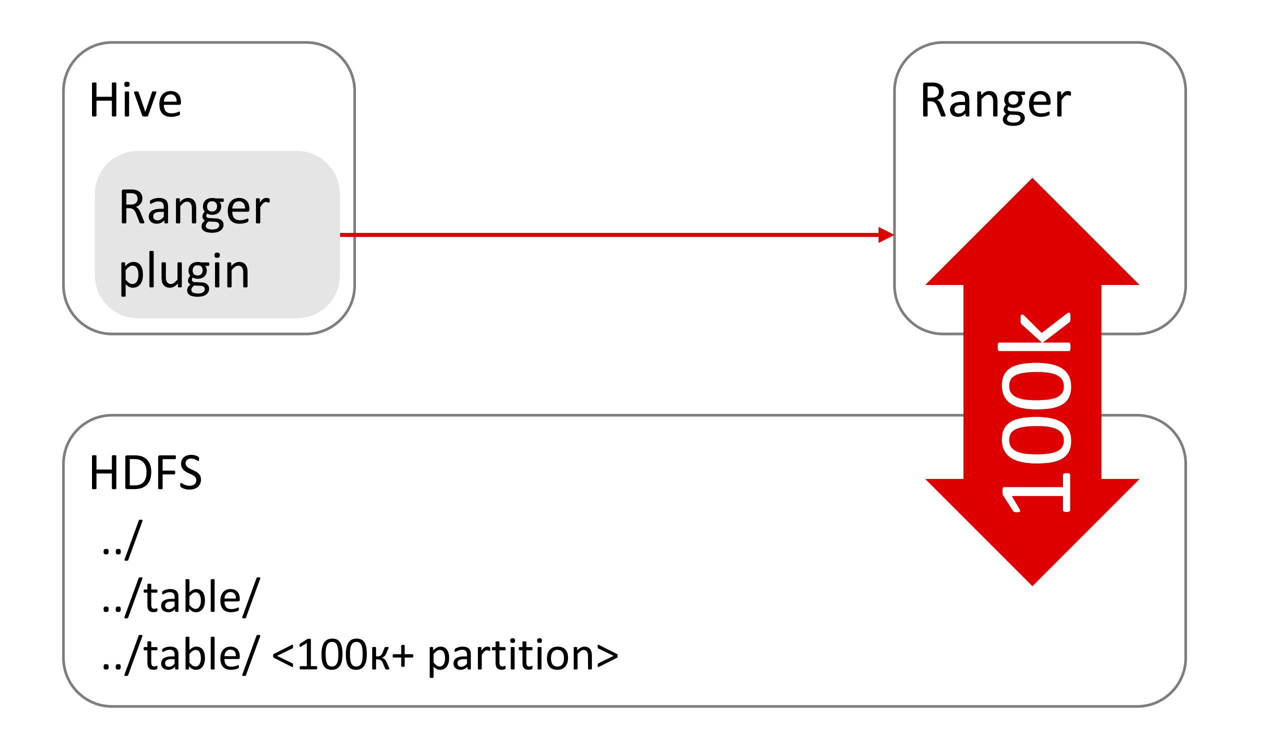
- Грануляция привилегий hdfs snapshot
- CVE-фикс для Spark 3
- Аудит Ambari
- Совместимость Ranger и Hbase





Доработки Hive

create external table big_table





ARM

VS

x86



Как определить что пора?

Потребности

Возможности

Сроки

Cost 2 value-анализ



Если бы пришлось повторить

2018	2019				2020		
Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3

PoT → Релиз 1.0 → Релиз 2.0 → Релиз 2.x

Team → Team → Team → Team

24 – 2 2 – 3 – 3 4 – 4 = 12



Какую команду собирать



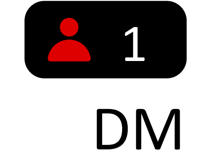








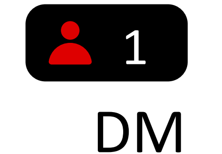


















Выводы

1

Верить в идею

2

Если делать, то получится

3

Есть Команда - есть результат



СПАСИБО

t.me/samorlov





Сцена после титров

1. Что будет с классическим дистрибутивом ближайшие 3 года

2. Как трансформируется платформа обработки данных в Сбере

3. Когда Hadoop действительно умрет или переродится

